Using Augmented Reality for Visualizing Complex Graphs in Three Dimensions

Daniel Belcher[1], Mark Billinghurst[1, 2], SE Hayes[1], and Randy Stiles[3]
[1]*Human Interface Technology Lab, University of Washington, Seattle*
[2]*Human Interface Technology Lab (New Zealand), University of Canterbury, Christchurch, NZ*
[3]*Lockheed Martin Corporation*
{[1]danb, [2]grof}@hitl.washington.edu, beephree@u.washington.edu, [3]randy.stiles@lmco.com

Abstract

In this paper we explore the effect of using Augmented Reality (AR) for three-dimensional graph link analysis. Two experiments were conducted. The first was designed to compare a tangible AR interface to a desktop-based interface. Different modes of viewing network graphs were presented using a variety of interfaces. The results of the first experiment shows that a tangible AR interface is well suited to link analysis. The second experiment was designed to test the effect of stereographic viewing on graph comprehension. The results show that stereographic viewing has little effect on comprehension and performance. These experiments add support to the work of Ware and Frank, whose studies showed that depth and motion cues provide huge gains in spatial comprehension and accuracy in link analysis.

1. Introduction

In two studies, Colin Ware's group showed that kinetic depth cues – and to a limited extent stereo depth cues – can greatly increase comprehension of complex graph structures. For comprehension, a 300% improvement in the size of the connected graph was observed, when compared to the same graph displayed in two dimensions [7]. This work builds on that of Sollenberger and Milgram [5][6] who found that individually stereo and kinetic cues produce comparable performance gains over two dimensional visual cues and that together they generate better performance that each one alone.

Subsequent work by Ware and Rose [8] has shown that the mode of manipulation is also a crucial element in the mapping between the input device and virtual object. Interfaces that use modes of manipulation that are closer to those found in everyday physical interaction, and that have simple spatial logic are what we call tangible interfaces. They find that tangible interfaces provide increased comprehension of graph structure.

Augmented Reality (AR) typically involves the overlay of virtual imagery on the real world. In this paper we explore how AR interfaces can be used to view complex linked graph structures. We believe that AR techniques could be beneficial for several reasons, including:

- Increased comprehension of complex link analysis graphs
- A large virtual display of graphical links in AR
- Enhanced spatial recall of link analysis graphs

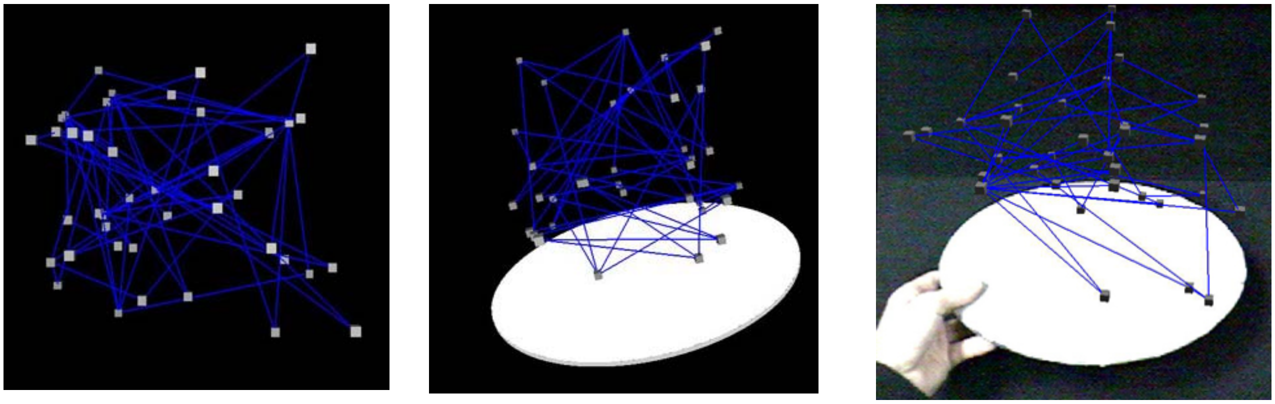
For many applications such as network analysis of web sites, circuits, and network routing, there is a need to understand connectivity between linked nodes displayed as a graph. This is often called link analysis. Ware and Frank showed that the interface could play a crucial role in accurate and rapid link analysis. Tangible interfaces provide a platform for building useful applications that have improved graph comprehension, but the question remains: does this effect transfer to Augmented Reality? We hypothesize that the same increase in comprehension for connected graphs occurs for Augmented Reality information displays as for screen-based 3D interfaces. The use of a tangible object for an input device may provide an additional positive effect. Finally, we hypothesize that an AR interface with a stereoscopic display system will perform as well as a similar 3D on-screen system using stereographic glasses.

To test these hypotheses we conducted two experiments. In the first, we were concerned with comparing a 2D on-screen viewing condition, to 3D on-screen and Augmented Reality viewing conditions. Participants were shown randomly generated graphs, at different levels of complexity, and asked questions about the connectivity of selected nodes. This study helped us determine the suitability of AR for link analysis.

In our second experiment, we were concerned with the effects of stereo on link analysis. Ware and Frank [7] showed little positive effect for stereo when compared with that of kinetic depth effects. Although we believe facility of movement and a simplified interface will provide the most gain, it is incumbent upon us to test stereo effects in our own interface. Therefore, our second experiment compared on-screen interfaces, both monoscopic and stereoscopic, to the corresponding mono- and stereo-AR interfaces.

Although several groups have used AR interfaces for scientific [2] and mathematical [4] visualization, as yet

Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '03)
0-7695-2006-5/03 $17.00 © 2003 **IEEE**



1: 2D on-screen condition **2: 3D on-screen condition** **3: AR condition**

Figure 1: Experiment 1 Conditions

there have been no user studies conducted comparing performance with these interfaces to screen-based systems. If our research shows that AR interfaces perform as well, or better than screen-based interfaces on graph analysis tasks then this may have significant implications for future AR visualisation interfaces.

2. Experiment 1: Modes of Viewing

The question we were asking in this experiment was "How different is an AR interface than a 2D – or 3D – on-screen interface for path tracing in a graph?" In order to get a quantitative evaluation, we have adopted (and adapted) the task used in Ware and Franck [7]: path tracing in an interconnected graph. The independent variable was the viewing condition. The dependent variables were percent error and response time.

Three viewing conditions (Figure 1) were employed:

1. *2D on-screen interface:* The 3D graph was shown on screen projected onto a 2D plane with a black background. An OpenGL projection matrix view was used. The viewer was allowed to zoom toward and away from the projection plane using the right mouse button. There was no rotation of the model.
2. *3D on-screen interface:* The 3D graph was shown on-screen atop a white disk placed to mimic the tracking disk in the AR condition. The subject used the left-mouse button to spin the viewpoint relative to the 3D graph in a trackball-like fashion. The right mouse button zoomed the viewpoint in and out.
3. *Augmented Reality interface:* The subject manipulated a real disk while wearing a Head Mounted Display with a small video camera attached to it. The disk had a square marker on it that was used by computer vision tracking software [1] to track the users viewpoint and overlay the graphic on the disk. The tracking marker was then overlaid with a smaller virtual disk that covered the marker itself. When the subject rotated or tilted the disk, the graph was rotated or tilted accordingly. The AR interface was also viewed against a black cloth background to ensure that the visual conditions were as close as possible to the 3D on-screen condition.

Note that both the 3D on-screen interface and AR interface were not showing stereo graphics. In the AR condition the application was running at 30 frames per second, the framerate of the video capture hardware, while in the other conditions it was running at 60 frames per second.

In previous work we have found that attaching virtual imagery to physcial objects have enabled users to manipulate virtual models in an extremely intuitive manner [3]. This ease of interaction should allow users to more effectively examine the node network. Thus we predict that the tangible AR interface will perform as well as the 3D on-screen interface, both in terms of accuracy (or minimal percent error) and response time.

2.1. Equipment

The hardware used in this study consisted of one Dell Dimension 4500 Pentium 4 with a PNY Nvidia Quadro4 XGL Series graphics card. For the Augmented Reality condition, an ELMO mini-camera was mounted on a SONY Glasstron LDI-D100B Head Mounted Display (SVGA 800x600, 38-degree diagonal field of view). A Hauppauge WinTV GO 190 video capture card was used to capture the NTSC signal from the video camera. In the AR condition the subjects manually manipulated a cardboard disk with a square tracking marker positioned

in the center (figure 2). In the two on-screen conditions, a two-button mouse was used to manipulate the orientation and position of the graph displayed on an SVGA monitor.

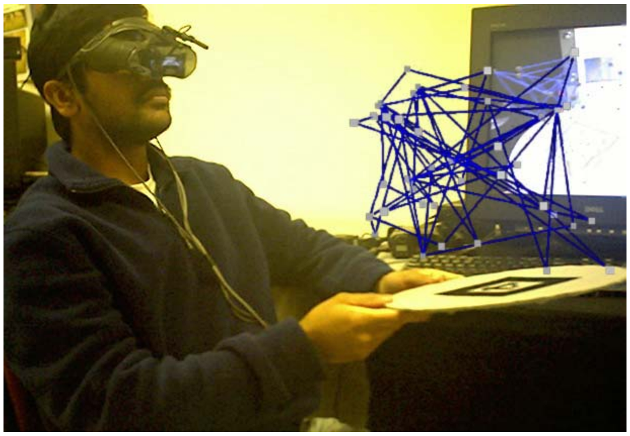


Figure 2: The AR Experimental Condition

2.2. Participants

Sixteen subjects took part in this experiment, ranging between 20 and 48 years old, with the average age being 25. The female to male ratio was 5:11. Two of the subjects were left-handed; one ambidextrous; and the remaining thirteen were right handed.

2.3. Experimental Procedure

The task was to decide whether there was a path consisting of only two connections between two highlighted nodes in a randomly generated graph. For each trial there was either a path of length two – from one highlighted node to another with only one intervening node between – or no such path (Figure 3).

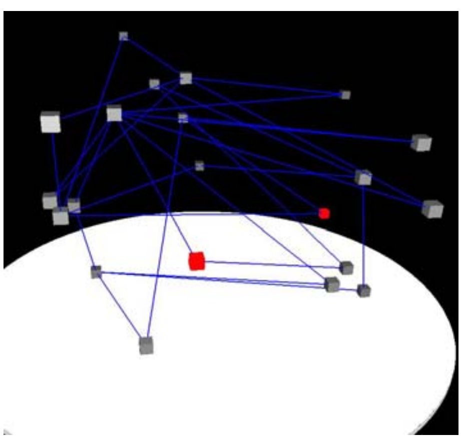


Figure 3: No Connected Path

The 3D graphs displayed consisted of a number of small nodes arranged randomly in a virtual volume (a cube 17cm on each side). In order to generate such graphs, the nodes were divided into three equally sized groups. Two such groups were the "potentially highlighted" nodes and the third group consisted of the "intermediate nodes." Each potentially highlighted node was connected via arcs to two different nodes in the intermediate group. For n nodes, this produced a total of (4/3 * n) connecting links. All the nodes were placed randomly in the simulated volume. Five levels of graph complexity were used:

Level	Num. of Nodes	Num. of Links
1	21	28
2	36	48
3	48	64
4	63	84
5	75	100

This resulted in 15 combinations of graph level and conditions. There were 10 trials per level and all were performed in each condition, resulting in 150 trials per experimental session.

Highlighted nodes were drawn in bright red, and the unhighlighted (intermediate) nodes were drawn in gray. Lighting was applied to all nodes in the graph. Each node was set to be 0.4 cm on each side. The connecting arcs were drawn in blue with two-pixel lines. The background color was black in all conditions (see Figure 3).

Upon arrival at the lab, the subjects were given a set of simple written instructions and were allowed two minutes to practice in each condition and ten practice trials. Subjects were instructed to "*answer as accurately and quickly as possible*" and to respond with a verbal "*yes*" or "*no*" as to whether the nodes were connected or not. The experimenter recorded the subject's answer by pressing either 'y' or 'n' on the keyboard. The orders of the conditions, as well as the order of the size of the graphs presented, were blocked and counterbalanced across all subjects. Prior to each trial, subjects were told which graph level to expect. During each trial, subject were given as much time as necessary to respond. The response time and response validity were recorded.

2.4. Results

The error data for this experiment is summarized in Figure 4. The x-axis represents the graph complexity, as reflected in the number of nodes. The y-axis represents the mean percent error.

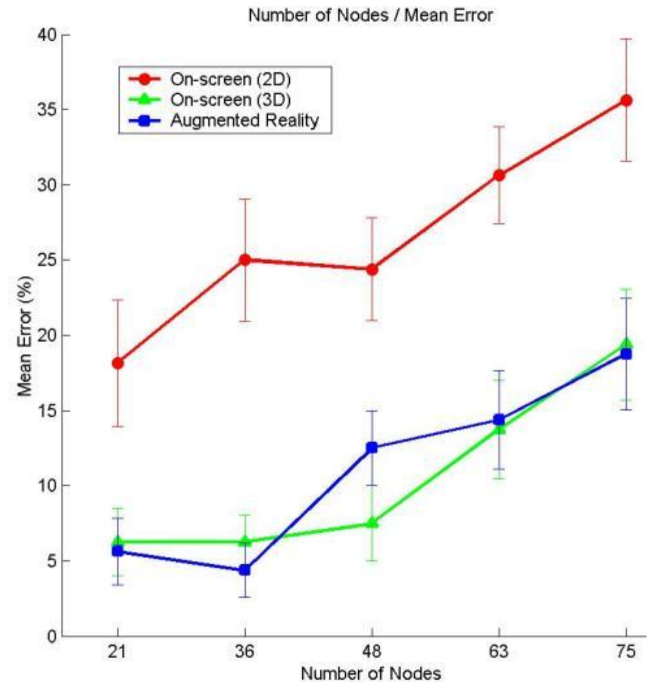


Figure 4: Mean Percent Error for Experiment 1

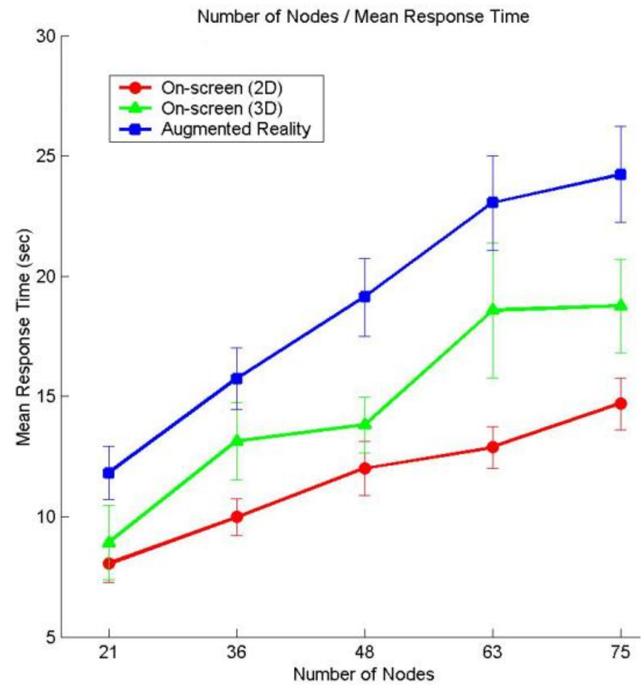


Figure 5: Mean Response Time for Experiment 1

As can be seen, the percentage of errors in the 2D on-screen condition is greater than that of the 3D on-screen condition and the AR condition. These results are virtually identical to, and consistent with, those found in [7]. An analysis of variance revealed a significant main effect of condition on mean percent error, $F(2,12) = 11.021$, $p < 0.05$. An ANOVA revealed a significant difference between the 3D and 2D conditions for mean error with $F(1,8) = 16.707$, $p < 0.05$ and a significant difference between 2D and AR, with $F(1,8) = 15.103$, $p < 0.05$. An ANOVA revealed no significant difference between the 3D and AR conditions.

Figure 5 summarizes the time data in this experiment. It shows a series of curves roughly separated by one standard deviation of error from each other. The x-axis represents the number of nodes. The y-axis represents the mean response time in seconds. Surprisingly, the AR condition took the most time to complete of the conditions. Furthermore, the 2D condition took the least amount of time across all levels of complexity. An analysis of variance revealed a significant main effect of condition on response time with $F(2,12) = 3.975$, $p < 0.05$ and a significant difference between the AR and 2D conditions, $F(1,8) = 7.986$, $p < 0.05$

2.5. User Feedback

After the experiment the subjects were asked a number of questions. First they were asked to "*rank the three interfaces in terms of overall ease of use*," with 1 being assigned to the easiest condition, and 3 to the most difficult. Figure 6 shows that users preferred the 3D on-screen and AR conditions over the 2D condition. An analysis of variance revealed a significant main effect of condition on ranking, $F_{(2,45)} = 7.34$, $p < 0.01$.

The second statement was: "*please rank the three interfaces/conditions in the order in which you preferred the information displayed,*" with 1 being the condition most preferred, and 3 being the least (see figure 6). An analysis of variance revealed a significant main effect of condition on the information display preference ranking, $F_{(2,21)} = 14.38$, $p < 0.01$.

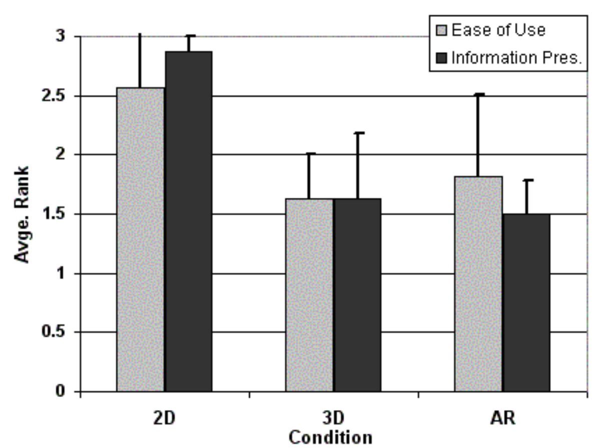


Fig. 6: Ease of Use and Information Preference

The next question was *"rank the three interfaces/ conditions in terms of ease of physical manipulation of the position and orientation of the graphic"* with a score of 1 being the easiest and 3 being the most difficult (see figure 7). An analysis of variance revealed no significant main effect of condition on the manipulation ranking.

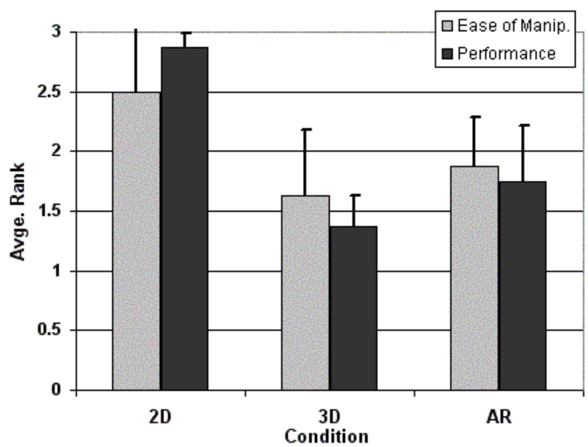


Figure 7: Avge. Rankings of Ease of Manipulation and How Well Task Performed

The final ranking was phrased as: "*rank the three interfaces in terms of how well you believe you performed the given task*" with 1 the best, and 3 the worst (see figure 7). An analysis of variance revealed a significant main effect of condition on perceived performance ranking, $F_{(2,45)} = 35.1$, $p < 0.01$. The majority of subjects believed they performed best in the 3D on-screen condition, and many participants thought they performed worst in the 2D on-screen condition.

2.6. Discussion

On the survey, 15 of the 16 respondents ranked the conditions identically for the *"overall ease of use"* and *"ease of physical manipulation"* scores. This may indicate the potential of AR lies in its usability and tangibility. This is supported by the rankings between the AR and the 3D condition, which are strikingly similar.

One of the most common complaints in the AR condition was the marker tracking. In order to have effective AR tracking the tracking marker needed to be in view at all times. This was not possible because a virtual disk covered the marker. Thus the virtual graph model sometimes flicked in and out of view.

Even though the users performed as accurately in the AR interfaces they took more time than in the 2D and 3D screen conditions. This may have been because of the perceptual qualities of the HMD compared the screen. Although the screen and HMD had the same resolution, they both had different fields of view and color and contrast properties.

Subjects were also asked to describe the strategy they used. The most common was a "process of elimination". Respondents would focus on one of the highlighted nodes and then trace a path to the immediately connecting nodes to see if their arcs angled back toward the other highlighted node. The most common variation was to look at the angle of the arcs leaving both highlighted nodes in an attempt to determine if there was a possible viewpoint from which these lines intersect.

3. Experiment 2: Stereo vs. Mono

Experiment 2 explored the effect of stereopsis. The task remains unchanged: path tracing in an interconnected graph. As in the previous experiment, the independent variable was the viewing condition. The dependent variables were percent error and response time. Four viewing conditions were employed:

1) On-screen Mono: This was the same as 3D on-screen interface in Experiment 1. However, the background was gray and a texture map of the tracking marker used in the AR condition was attached to the virtual tracking disk.

2) On-screen Stereo: Same as condition 1, except in stereo. The correct view of the graph was generated for each eye position and continuously updated. The subject wore a pair of LCD shutter glasses to view the screen image in stereo.

3) Augmented Reality Mono: This interface was the same as the AR interface in Experiment 1, with the tracking marker no longer occluded.

4) Augmented Reality Stereo: Same as the preceding condition, except in stereo. Two cameras (instead of one) were mounted on the HMD and the AR image was presented using quad-buffered stereo video with graphic overlay.

Using the findings of Ware and Frank [7], we predict that stereo will add little positive effect to performance, both in terms of accuracy and response time.

3.1. Equipment

As in Experiment 1, the experimental display system for the AR conditions consisted of a SONY Glasstron LDI-100B HMD. However, in this experiment, two ELMO mini-NTSC cameras were mounted on the HMD to provide stereo video. A second video capture card was installed to handle the second video stream. The two cameras were mounted 5.8 cm apart (the average inter-ocular distance) on the top of the HMD, at such an angle (1.5 degrees) as to focus on an area roughly 85 cm away

(slightly less than the average arm length). The Sony Glasstron LDI-100B supports quad-buffered stereo. The same tracking card and pattern were used.

The StereoGraphics CrystalEyes 3D LCD shutter glasses, in synch with the emitter and the monitor provided the stereo effect in the on-screen Stereo condition. The vertical refresh rate of the monitor was set to 120 Hz, with each eye receiving an update at 60 Hz (Figure 8). In this case the virtual graphics were drawn with the same eye separation as in the AR condition.

Figure 8: The on-screen stereo setup.

3.2. Participants

Participants in Experiment 2 consisted of 16 subjects, ranging between 18 and 48 years old, with an average age of 25. The female to male ratio was 8:8. Only one of the subjects was left-handed; the remaining 15 were right handed. One of the subjects had previously used an AR interface, but none had previously used a stereographic display system.

3.3. Experimental Procedure

The experimental procedure in Experiment 2 was virtually unchanged from that of the preceding experiment with the same task. However, due to the increased number of conditions, and in order to complete the required number of trials in the allotted experimental time-slot, we were forced to lower the number of levels of graph complexity from 5 to 4. The range of complexity – from 21 to 75 – remained the same. The levels of graph complexity were as follows:

Level	Num. of Nodes	Num. of Links
1	21	28
2	39	52
3	57	76
4	75	100

This resulted in 16 graph level to condition (interface) combinations. There were 10 trials per level and all levels were performed with each interface. This resulted in a total of 160 trials per experimental session.

The color and lighting of the nodes and arcs remained unchanged from Experiment 1. However, in order to reduce the ghosting affects associated with stereo, the background color was changed to gray in all conditions. Subjects wore the shutter glasses in both the on-screen stereo and on-screen mono conditions. The tracking marker in the AR conditions was also no longer occluded and a texture map of the square and symbol was placed in the on-screen conditions. The decision to do this was at the request of Experiment 1's subjects (see section 2.5).

Upon arrival at the lab, the subjects were given a set of simple written instructions outlining the task. Before beginning the experiment proper, subjects were allowed two minutes to practice in each condition and ten practice trials to make sure they understood the task. Subjects were instructed to "*answer as accurately and quickly as possible*" and to respond with a verbal "*yes*" or "*no*". The experimenter recorded the subject's answer by pressing either 'y' or 'n' on the keyboard.

The order of the conditions, as well as the order of the size of the graphs presented, were blocked and counterbalanced across all subjects. Prior to each block of trials, subjects were told which graph level to expect. During each trials, subject were given as much time as necessary to respond. The response time and response validity were recorded. All participants were given as many breaks as they required and the overall experiment duration averaged one hour and fifteen minutes.

3.4. Results

The error data for this experiment is summarized in Figure 9 overleaf. The x-axis represents the graph complexity, as reflected in the number of nodes. The y-axis represents the mean percent error. As expected the mean error increases with the number of nodes, but there is little difference in error percentage across conditions for a given graph complexity. An analysis of variance for this data reveals no significant main effects or differences between conditions.

Figure 10 summarizes the time data in Experiment 2. It shows a series of curves almost indistinguishable. The x-axis represents the number of nodes, from 21 to 75. The y-axis represents the mean response time as measured in seconds. An analysis of variance for this data reveals no significant effects or difference between conditions.

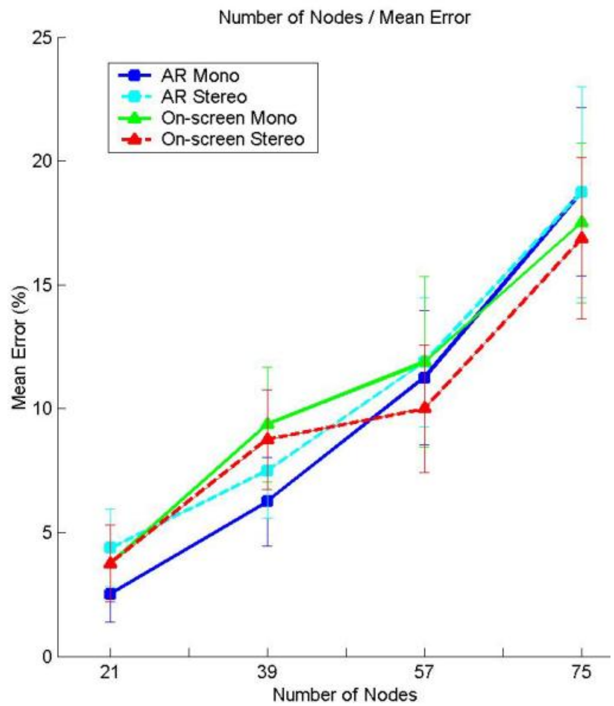


Figure 9: Mean Percent Error for Expt 2.

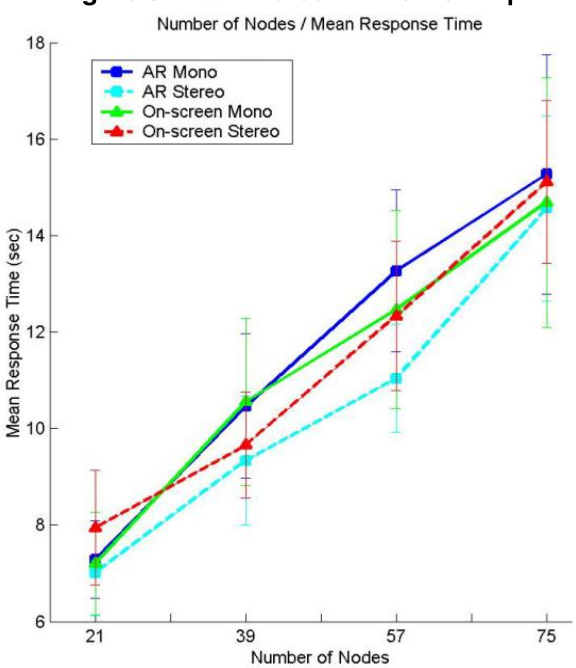


Figure 10: Mean Response Time for Expt 2.

3.5. User Feedback

As in Experiment 1, the participants were asked for their feedback on the experimental conditions. The quantitative results for this survey consisted of two ranking questions. In the first ranking, subjects were asked to "*rank the four interfaces in terms of overall ease of use*," with 1 being the easiest, and 4 the most difficult. The responses are summarized in Figure 11. The four conditions tested are on the x-axis, while the average ranking are displayed on the y-axis.

Figure 11 shows that subjects felt there was little difference in ease of use among the four conditions. It should be noted that those subjects who did not notice any particular difference between the AR Mono and the AR Stereo or between the on-screen Mono and the on-screen Stereo conditions decided to evenly rank the two respective conditions. So a subject who did not differentiate between AR Stereo and AR Mono simply gave the same ranking to both. This resulted in more than first (1) and second (2) place rankings for each condition that third (3) or fourth (4) rankings. An analysis of variance revealed no significant main effect of condition on ranking.

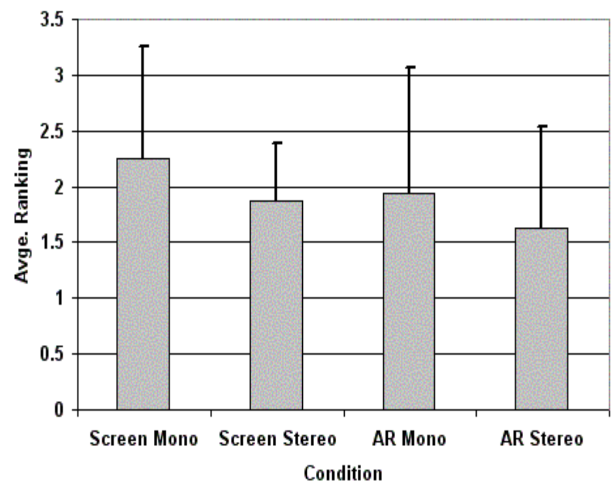


Figure 11: Avg. Rankings of Ease of Use

The second question was: "*please rank the four interfaces in terms of how well you believe you performed the given task*" with 1 being the best performance, and 4 being the worst performance. The responses to this ranking are summarized in Figure 12 overleaf.

Figure 12 shows that the many of subjects believed they performed best in the AR stereo condition. However, an analysis of variance revealed no significant main effect of condition on ranking. It should be noted that those subjects who did not notice any particular difference between the AR Mono and the AR Stereo or between the on-screen Mono and the on-screen Stereo conditions decided to evenly rank the two respective conditions. So a subject who did not differentiate between AR Stereo and AR Mono simply gave the same ranking to both. Subjects that reported ranking the AR

Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '03)
0-7695-2006-5/03 $17.00 © 2003 **IEEE**

Stereo condition as their best performance (with a ranking of 1) claimed to have done so because "they could better distinguish between the lines (arcs)."

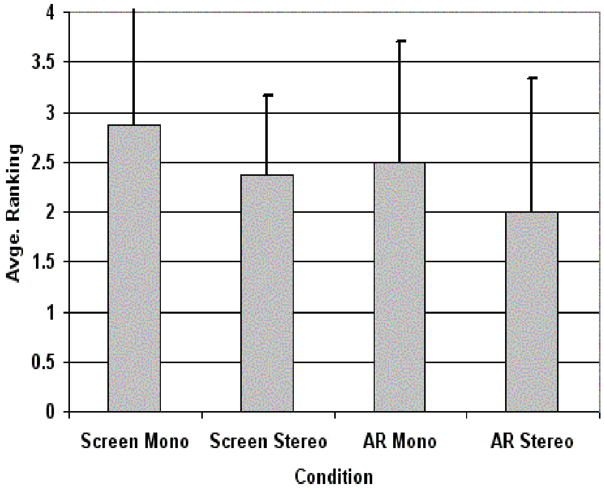


Fig 12: Avg. Rankings of Perceived Performance

3.6. Discussion

The search strategies used were virtually identical to those reported in Experiment 1. Most of the subjects reported using the "process of elimination" strategy, after a small amount of exploration. Many of the subjects (10 of 16) reported to have "just looked for the intermediary node" in the on-screen Stereo and AR Stereo conditions in which the node count was low (21 or 39).

One of the limitations of stereo video-capture based Augmented Reality is the fixed camera angles. In our configuration the focal length of the cameras remains fixed no matter how near or far the objects in the field-of-view. This becomes a problem when the user wishes to zoom in close to the tracking-card, as the camera angles do not change with the distance of the object. This limitation is not inherent in screen-based quad-buffered stereo (on-screen stereo condition) because the OpenGL buffer rendering code dynamically changes the focal length based upon the location of the 3D object.

4. Conclusion

In this paper we reproduced the classic experiments of Ware and Franck [7] to evaluate the usefulness of using AR for visualizing complex three-dimensional node-graph representations.

In this paper we have performed two related experiments. In the first we compared link understanding in 2D and 3D screen conditions to an AR interface. Subjects took longer in the AR condition than the two screen-based conditions, but produced as few errors as the 3D screen condition and significantly less than the 2D screen case. In subjective rankings of Ease of Use, Information Display Preference, Ease of Manipulation and Perceived Performance, subjects felt that the AR condition was equivalent to the 3D screen condition and significantly better than the 2D screen case.

The second experiment explored the effect of adding stereo cues and compared accuracy and timing results across four conditions (stereo and non-stereo screen interfaces and stereo and non-stereo AR). Although performance got worse as the number of graph nodes increased, there was no difference in performance between these conditions at a given graph complexity.

The major difference between the 2D screen condition and the 3D and AR conditions is in the ability to rotate the model. Thus these results support those of Ware and Franck, namely that graph understanding is significantly improved by the support for kinetic depth cues. The lack of performance and accuracy differences between the stereo and non-stereo conditions similarly highlight the dominance of kinetic over stereo cues for this task.

In experiment one, users performed as accurately with the AR interface as with the 3D screen interface and felt it was just as good to use. This implies that AR interfaces could be an effective way to visualize abstract information such as interconnected graphs. One key advantage of AR interfaces is the support for a tangible interaction metaphor, with its direct mapping between a real-world physical object and virtual object. This enables users to easily manipulate the graph content and view it from any perspective, improving their comprehension.

Although these results are interesting, they are just the beginning. In the future we want to explore other types of graph visualization and comprehension tasks and compare between screen-based and AR visualization. We also want to explore a wider range of manipulation and interaction techniques for AR interfaces for scientific visualization.

5. Acknowledgements

This work was supported by the Advanced Research and Development Activity (ARDA) under contract number NMA201-01-C-0006. Thanks to Hannah Slay of the WCL at the University of South Australia for her help and insight. Thanks to Kiyoshi Kiyokawa of Osaka University for his help, criticism and encouragement. We also thank Konrad Schroder of the HIT Lab Seattle for much needed technical support. Finally, thanks go to the lab members at the HIT Lab Seattle and the HIT Lab New Zealand for providing such an amazing collaborative atmosphere.

6. References

[1] ARToolKit 2001. ARToolKit website: http://www.hitl.washington.edu/artoolkit.

[2] A. Fuhrmann, H. Löffelmann, D. Schmalstieg "Collaborative Augmented Reality: Exploring Dynamical Systems" *IEEE Visualization 1997*, pp. 459-462, November 1997.

[3] Kato, H., Billinghurst, M., Poupyrev, I., Tetsutani, N., Tachibana, K. Tangible Augmented Reality for Human Computer Interaction (in Japanese) *The Journal of the Society for Art and Science,* vol.1, no.2, pp.97-104, 2002.

[4] H. Kaufmann, D. Schmalstieg, M. Wagner. "Construct3D: A Virtual Reality Application for Mathematics and Geometry Education" *Education and Information Technologies* 5:4, special issue on "Virtual Reality", pp. 263-276, 2000.

[5] R.L. Sollenberger, and P. Milgram. "A comparative Study of Rotational and Stereoscopic Computer Graphic Depth Cues." *Proceedings of the Human Factors Society Annual Meeting*, 1452-1456, 1991.

[6] R.L. Sollenberger, and P. Milgram. "The effects of Stereoscopic and Rotational Displays in a Three-Dimensional Path-Tracing Task." *Human Factors*, 35 (3) 483-500, 1993.

[7] C. Ware, and G. Franck. "Evaluating Stereo and Motion Cues for Visualizing Information Nets in Three Dimensions." *ACM Transactions on Graphics*, 15(2) 121-139, 1996.

[8] C. Ware, and J. Rose. "Rotating Virtual Objects with Real Handles." *ACM Transactions on CHI*, 6(2) 162-180, 1999.

0-7695-2006-5/03 $17.00 © 2003 **IEEE**